Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 34 (2018) 163–170

Transportation Research Procedia

www.elsevier.com/locate/procedia

International Symposium of Transport Simulation (ISTS'18) and the International Workshop on Traffic Data Collection and its Standardization (IWTDCS'18)

A Lane-based Predictive Model of Downstream Arrival Rates in a Queue Estimation Model Using a Long Short-Term Memory Network

Seunghyeon Lee[1], Kun Xie[1]*, Dong Ngoduy[1], Mehdi Keyvan-Ekbatani[1], Hong Yang[2]

[1] *Department of Civil and Natural Reources Engineering, University of Canterbury, Christchurch, New Zealand*
[2] *Department of Modeling, Simulation & Visualization Engineering, Old Dominion University, Norfolk, USA*

Abstract

In this study, we develop a mathematical framework to predict cycle-based queued vehicles at each individual lane using a deep learning method – the long short-term memory (LSTM) network. The key challenges are to decide the existence of residual queued vehicles at the end of each cycle, and to predict the lane-based downstream arrivals to calculate vertical queue lengths at individual lanes using an integrated deep learning method. The primary contribution of the proposed method is to enhance the predictive accuracy of lane-based queue lengths in the future cycles using the historical queuing patterns. A major advantage of implementing an integrated deep learning process compared to the previously Kalman-filter-based queue estimation approach (Lee et al., 2015) is that there is no need to calibrate the co-variance matrix and tune the gain values (parameters) of the estimator. In the simulation results, the proposed method perform better in only straight movements and a shared lane with left turning movements.

Keywords: Type your keywords here, separated by semicolons ;

1. Introduction

The estimation method of queue lengths has been playing a significant role in traffic management and

* Corresponding author. Tel.: +64-3-369-2707; fax: +64-3-364-2758.
E-mail address: kun.xie@canterbury.ac.nz

2352-1465 © 2018 The Authors. Published by Elsevier Ltd.

10.1016/j.trpro.2018.11.028

control. Moreover, the number of dynamic queued vehicles is needed to construct the adaptive signal control logic in real time. Webster (1958) and Akcelik (1980) introduced the fundamental concepts of queue lengths and delay, which are the cornerstones for the development of diverse methods for estimating queue lengths, involving the conservation equation (Lindley, 1952), the shockwave theory (Lighthill and Whitham, 1955; Richards, 1956) and the probe vehicle approach (Comert and Cetin, 2009; Cheng et al., 2011, 2012; Christofa et al., 2013).

We focus on the improvement of methods based on the conservation equation, which have been widely adopted for the estimation of queue lengths over the past several decades due to their simplicity, robustness, and applicability. Lee et al. (2015) develop the improved conservation equation including discriminant models based on logistic regression models and estimation methods of downstream arrivals using the Kalman filters. The role of discriminant models is to identify the existence of residual queues in each lane at the start time of every cycle. The cumulative errors induced by passage detectors or counting errors are eliminated by the discriminant models. The average time occupancy rate at downstream detectors per each lane, the traffic flows for red and green period during a cycle, and the average occupancy rate at upstream detectors per each approach have a significant influence on the probability of the existence of an initial queue in an individual lane. Furthermore, the proportion of upstream arrivals and discharges per each lane are defined to estimate lane-changing proportions in an area between the upstream and the downstream detectors for downstream arrivals. Although the modified conservation equation shows improved estimates of lane-based queued vehicles in real time compared with the existing methods in Lee et al. (2015), there is still the potential for improvement of estimation procedures of both an identification of a residual queue in discriminant models and a proportion of downstream arrival in the Kalman filters. The developed queue estimation model is used in a predictive model of lane-based control delay (Lee and Wong, 2017b), derivatives of the control delay (Lee and Wong, 2017a), and group-based adaptive traffic-signal control (Lee et al., 2017a, 2017b).

In this study, we apply a deep learning method to estimate cycle-based proportions of downstream arrivals used to predict downstream arrivals for the number of vertically queued vehicles in the current cycle. The key challenges are to construct Long Short-Term Memory (LSTM) network, which is one of widely used Recurrent Neural Network (RNN) methods to solve time-series forecasting problems in LeCun and Bengio, 1995. The primary contribution of the proposed method is to enhance the predictive accuracy of lane-based queue lengths in the future cycles using the historical patterns of proportional downstream arrivals. A major advantage of implementing an integrated deep learning process compared to the previously Kalman-filter-based queue estimation approach (Lee et al., 2015) is that there is no need to calibrate the co-variance matrix and tune the gain values (parameters) of the estimator. The proposed method consists of discriminant models and a downstream arrivals estimation method using a comprehensive deep learning method to predict short-term queue lengths at each lane. The data processing method for the real-time predictive model is illustrated in Figure 1.

In Figure 1, occupancy rates and impulse memories at both upstream and downstream detectors are collected in seconds. Discriminant models are used to identify the residual queue length at the end of each cycle. The deep learning method is used to predict lane-to-lane proportions of traffic flows. In addition to the cycle-based recursive process, the real-time collected data is utilized to estimate downstream arrivals in a second using the improved conservation equations developed in Lee et al. (2015). Consequently, lane-based queue lengths are predicted applying both the cycle-based recursive predictive model and the real-time estimation method of downstream arrivals.

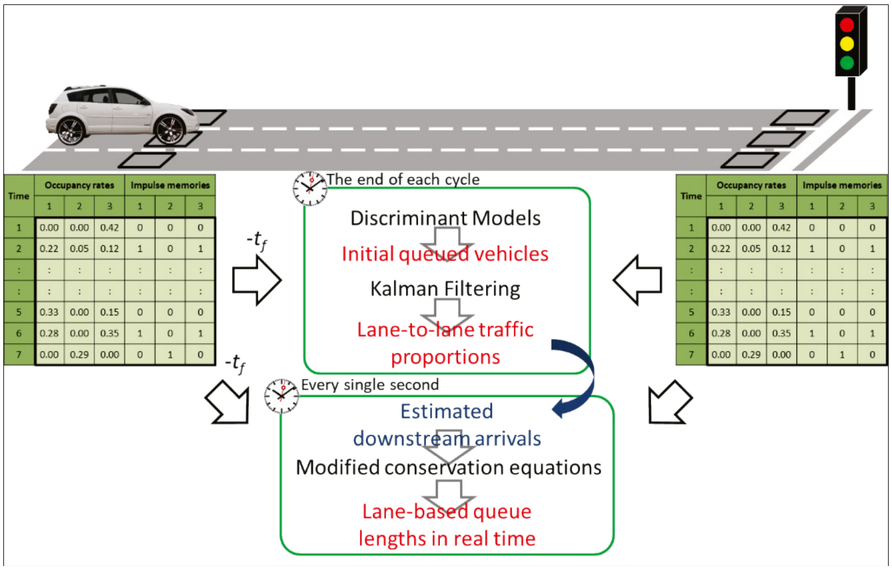

Fig. 1. Data-processing method for the deep learning approach to dynamic queue estimation.

2. Methodology

We develop a real-time estimation approach for lane-based queue lengths. For the discriminant models, we use logistic regression models to decide whether there are residual queues or not at the end of each cycle. The main role of the discriminant models is to prevent the cumulative errors induced by passage detectors or counting errors. In the logistic regression models, a binary variable has been defined to identify the existence of the residual queue (i.e. 1 and 0 indicate the presence and the absence of the residual queue, respectively). The first independent variable is the average time-occupancy rate collected by the downstream detector during the effective time duration. The traffic flows for red and green periods are set separately since the different levels of effectiveness in traffic flows for the red and green periods are considered in terms of utility. In these two variables, the proportions of downstream discharges in each lane are used to examine the lane-changing behaviour. The second and third variables are the traffic flows during the red and green periods at the upstream detectors. The last variable is the average time-occupancy rate, which is collected from the upstream detectors during a cycle.

The binary indicator of the discriminant models is used in the enhanced conservation equation (Lee et al., 2015). The traditional conservation equation is regarded as the benchmark for the dynamic queue lengths estimation method. To account for the lane-changing behaviour and the counting errors, the figure for estimated downstream arrivals is used to replace the upstream arrivals in the conventional equation. The proportion of lane-changing traffic flows is based on the proportion of lane-to-lane traffic volumes in the previous cycle using a series of equations in Lee et al. (2015). To predict the proportion of lane-changing traffic flows, we propose a novel deep learning method based on LSTM network. Below we list the indices, parameters, and variables used in our analysis:

Nomenclature

K	number of lanes
N	number of cycles
T	number of time steps in a cycle
t_f	the free-flow travel time from the upstream detector to the downstream detector
$o^n_{u,k}(t - t_f)$	the time occupancy from the upstream detector in lane k at time $t - t_f$ in the nth cycle
$o^n_{d,k}(t)$	the time occupancy from the downstream detector in lane k at time t in the nth cycle

$i^n_{u,k}(t-t_f)$	the upstream arrivals according to the binary impulse memory from the upstream detector in lane k at time $t-t_f$ in the nth cycle
$i^n_{d,k}(t)$	the downstream discharges according to the binary impulse memory from the downstream detector in lane k at time t in the nth cycle
$\hat{a}^n_{d,k}(t)$	the estimated downstream arrivals according to the estimated arrival rates from the downstream detector in lane k at time t in the nth cycle

2.1. Lane-based queue lengths predictive model

To prevent the cumulative errors generated by the traditional conservation equation, we use a discriminant model to determine whether there are any residual queues at the start time of each cycle. The discriminant model is a logistic regression model, and is specified as follows.

$$\mathrm{Prob}\{x_k^n(t)=1\}=\frac{1}{1+\exp(-u_k^n(t))} \tag{1}$$

where for each lane, the logit (or utility function) is defined as

$$\begin{aligned} u_k^n(t) &= \alpha+\beta_1\frac{1}{m}\sum_{t=T-m+1}^{T} o_{\mathrm{d},k}^{n-1}(t)+\beta_2\sum_{k=1}^{K}\sum_{t=1}^{R} i_{\mathrm{u},k}^{n-1}(t-t_{\mathrm{f}})\times p_{\mathrm{d},k}^{n-1}+\beta_3\sum_{k=1}^{K}\sum_{t=T-R}^{T} i_{\mathrm{u},k}^{n-1}(t-t_{\mathrm{f}})\times p_{\mathrm{d},k}^{n-1} \\ &+\beta_4\frac{1}{K}\frac{1}{T}\sum_{k=1}^{K}\sum_{t=1}^{T} o_{\mathrm{u},k}^{n-1}(t-t_{\mathrm{f}}) \end{aligned} \tag{2}$$

$$p_{\mathrm{d},k}^{n-1}=\frac{\sum_{t=1}^{T} i_{\mathrm{d},k}^{n-1}(t)}{\sum_{k=1}^{K}\sum_{t=1}^{T} i_{\mathrm{d},k}^{n-1}(t)} \tag{3}$$

The detailed explanations are provided in Lee et al. (2015). To apply the model, we determine the logits $u_k^n(t)$ and $\forall k\in K$, for the set of lanes using (2). We decide that there is a probability higher than 50% that there will be a residual queue on lane k at the start time of the nth cycle and that there is a probability less than 50% that a residual queue will becomes zero on lane k at the start time of nth cycle , i.e., we set the queue on each lane at the start time of the present cycle as equal to the queue on each lane at the end time of the previous cycle, such that $\mathrm{Prob}\{x_k^n(t)=1\}$ is higher than 50%.

Because upstream arrivals can be different than downstream arrivals on the same lane, for a real-time estimation of lane-based queue lengths the number of downstream arrivals on lane k at time t on the nth cycle, i.e., $\hat{a}_{d,k}^n(t)$, is required to estimate lane-to-lane queue lengths with high accuracy. The discordance between upstream and downstream arrivals on the same lane can be primarily attributed to lane-changing behavior and counting errors. Therefore, we improve the traditional conservation by using $\hat{a}_{dw,k}^n(t)$ to estimate downstream arrivals. The modified conservation equation is as follows.

$$q_k^n(t)=\begin{cases}\max\left\{x_k^n(t)q_k^{n-1}(T)+\hat{a}_{\mathrm{d},k}^n(t)-i_{\mathrm{d},k}^n(t),0\right\}, & \text{for } t=1\\ \max\left\{q_k^n(t-1)+\hat{a}_{\mathrm{d},k}^n(t)-i_{\mathrm{d},k}^n(t),0\right\}, & \text{for } t\neq 1\end{cases} \tag{4}$$

We basically use the proportion of the total downstream discharges on each lane and the proportion of the total traffic volume on each lane to calculate each lane's upstream arrivals for $\hat{a}^n_{d,k}(t)$. The proportions for $\hat{a}^n_{d,k}(t)$ are calculated based on data collected from upstream and downstream detectors during the n-1th cycle. We set the proportions of traffic volumes on lane k as

$$\hat{a}^n_{\mathrm{d},k}(t) = \sum_{k=0}^{K} i^n_{\mathrm{u},k}(t - t_{\mathrm{f}}) \times d^{n-1}_{\mathrm{d},k}, \tag{5}$$

$$d^n_{\mathrm{d},k} = \frac{\sum_{t=1}^{T} i^n_{\mathrm{d},k}(t)}{\sum_{k=1}^{K}\sum_{t=1}^{T} i^n_{\mathrm{d},k}(t)} \tag{6}$$

Equation (6) calculates the proportion of downstream discharges on lane k during the nth cycle and the proportion of upstream arrivals on lane k during the n-1th cycle given the free flow travel time. Simple growth factors are calculated using equations (5). Equation (5) compares the proportions of downstream discharges on lane k collected during the n-1th cycle to upstream arrivals. In this process, we directly use the proportions of upstream arrivals and downstream discharges data collected during the n-1th cycle to obtain $\hat{a}^n_{d,k}(t)$. To improve the accuracy of the estimates, in the next step we apply a recursive process, specifically the Kalman filter, to estimate downstream arrivals.

2.2. LSTM network for proportional downstream arrivals

The LSTM network is a special type of recurrent neural networks. It packages the hidden layers as memory cells and can account for the autocorrelation within time series data in both long and short-term by controlling the transfer of cell states and hidden states between different cells. To start the recursive process, the proportion of upstream arrivals in the previous and the current cycle, $d^{n-1}_{\mathrm{u},k}$ and $d^n_{\mathrm{u},k}$, and downstream discharges, $d^{n-1}_{\mathrm{d},k}$ are used as the input, meanwhile the proportion of downstream arrivals, $\hat{a}^n_{\mathrm{d},k}$, is defined as the output for cycle n. The set of $d^{n-1}_{\mathrm{u},k}$, $d^n_{\mathrm{u},k}$, and $d^{n-1}_{\mathrm{d},k}$ are used as an input, x_t, whereas $\hat{a}^n_{\mathrm{d},k}$ is used as an output, z_t, at step t in the LSTM network. The LSTM network updates for time step t given inputs x_t, h_{t-1}, and c_{t-1} are:

$$\begin{aligned} i_t &= \sigma\left(W_{xt}x_t + W_{hi}h_{t-1} + b_i\right) \\ f_t &= \sigma\left(W_{xf}x_t + W_{hf}h_{t-1} + b_f\right) \\ o_t &= \sigma\left(W_{xo}x_t + W_{ho}h_{t-1} + b_o\right) \\ g_t &= \sigma\left(W_{xc}x_t + W_{hc}h_{t-1} + b_c\right) \\ c_t &= f_t \odot c_{t-1} + i_t \odot g_t \\ ht &= o_t \odot \phi\left(c_t\right) \end{aligned} . \tag{7}$$

In Equations (7), h_t, i_t, f_t, o_t, g_t, and c_t are hidden units, input gates, forget gates, output gates, input modulation gates, and memory cells, respectively. $\sigma(x)$ and $\phi(x)$ are element-wise non-linearities, such as a sigmoid or hyperbolic tangent. Ws are the sequential model's weights and bs are the sequential model's biases. $\odot$ denotes the Hadamard product.

3. Model results

A typical directional link at an isolated intersection is created using AIMSUN (Barcelo, J. et al., 1989), which is widely used in microscopic simulations. The impulse memories and time occupancy rates of detectors can be collected using a function of a detection in AIMSUN. The geometric layout of the junction used in the simulation is shown in Figure 2. The cycle length is 60 seconds. There are four phases, one for each direction. The duration of the green signal is ten seconds for each phase. The inter-green time between any two consecutive phases is 5 seconds. To collect the impulse memory and time occupancy rate data at the upstream and downstream, we mimic a full set of 2 m × 2 m loop detectors. Upstream detectors are setback at 200 m from the stop-line and the free flow travel time is 12 seconds. We set the four levels of traffic volume on all of the approaches to 300, 600, 900, and 1200 vehicles per hour (vph) to train the LSTM network. The turning proportions are randomly generated among all of the movements. Right turns on red (RTOR) are not allowed in the simulation, as the driving rule is right-hand traffic. The simulation period is 24 hours and the simulation resolution is one second. The first 965 cycles are used for training the proposed model, whereas the last 473 cycles are used to validate the performance of the proposed model.

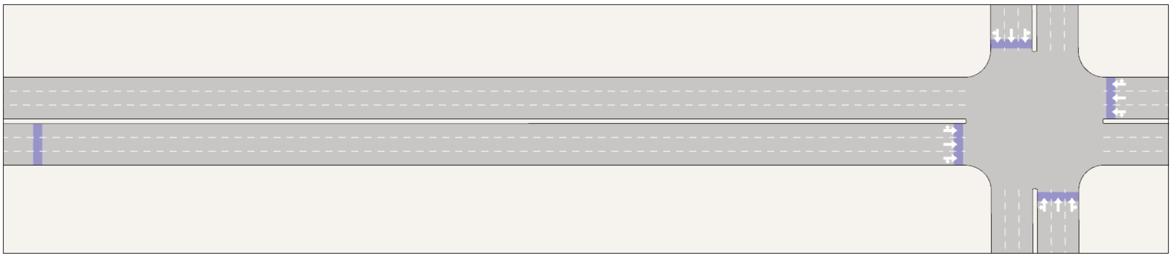

Fig. 2. Geometric Layout of the Simulated Junction in AIMSUN.

For a better appraisal of the prediction accuracy, the performance of the proposed model has been compared with three other lane-based queue estimation models. Model I estimates queue lengths based on the traditional conservation equation, which is commonly used to develop adaptive control logic models. Model II is the recently developed queue predictive model based on the Kalman filter (Lee et al., 2015). Model III, which is the suggested model of this manuscript, includes the calibrated discriminant models, the proportions of the aggregated lane-to-lane traffic volumes in each lane, and the predicted downstream arrivals predicted by the LSTM network. The RMSE values (veh • sec) per each lane of Model I-III to estimate lane-based queue lengths for 24 hours were specified in Table 1.

Table 1. RMSE values of the 24 hours queue estimation including train and test set.

	Model I			Model II			**Model III**		
	Lane 1	Lane 2	Lane 3	Lane 1	Lane 2	Lane 3	Lane 1	Lane 2	Lane 3
Train Set	12.093	12.671	31.302	12.185	6.480	12.356	13.885	**5.890**	**11.933**
Test Set	10.722	19.935	32.472	10.170	8.586	7.731	11.842	**7.931**	**7.566**

In Table 1, the RMSE values of the proposed model on lane 2 and lane 3 in train and test sets, 5.890, 11.933, 7.931, and 7.566 (veh • sec), respectively, are slightly lower than that values of the authors' previous model, 6.480, 12.356, 8.586, and 7.731 (veh • sec), respectively. Furthermore, our proposed model, Model III, shows overwhelming performance at lanes 2 and 3, compared with the traditional queue length estimation model, Model I. Although the performance of the Model II, which is our previous model, based on the Kalman filters, is slightly better than the Model III at lane 1 in both train and test sets, the accuracy of the Model III is higher than Model II at lane 2 and 3 in both train and test sets. The model III that is based on the discriminant model and included lane-changing behavior based on the LSTM network shows noticeably improved accuracy. We therefore further analyzed the accuracy fluctuations of these models as the estimated cycle-by-cycle average queue lengths. As illustrated in Figure 3 (a) – (c), we compare the average queue length per each cycle, estimated by the Model III, with that of observed queue lengths.

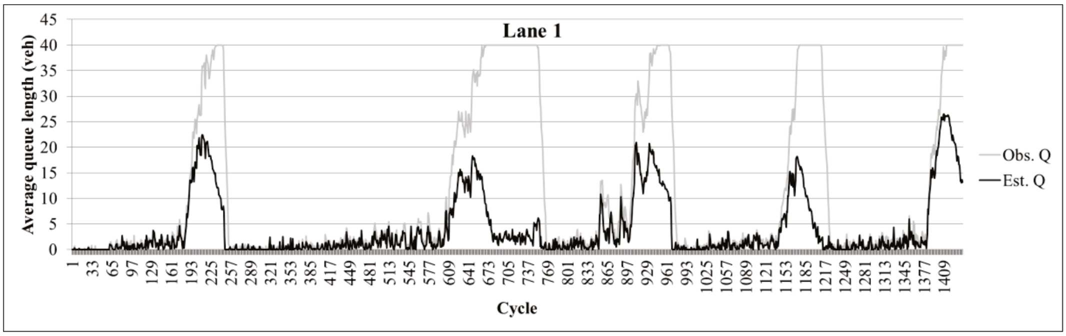

Fig. 3 (a). Comparison of average queue lengths as observed and estimated by Model III on lane 1.

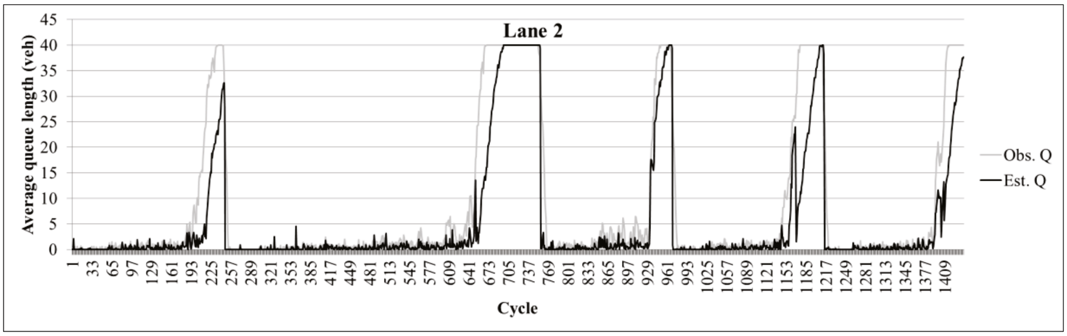

Fig. 3 (b). Comparison of average queue lengths as observed and estimated by Model III on lane 2.

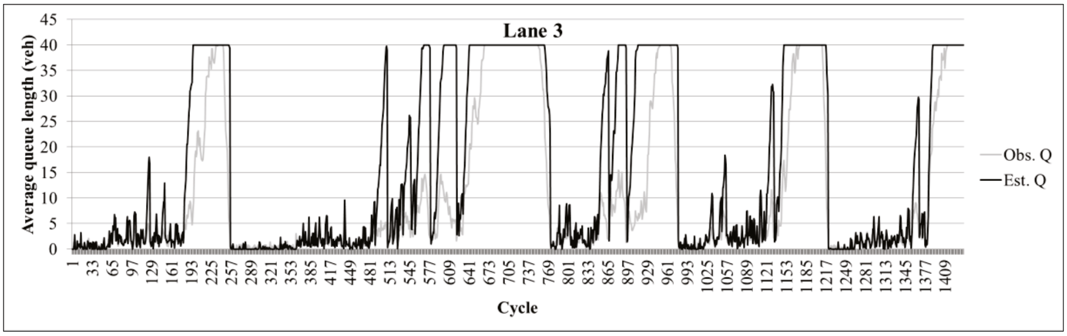

Fig. 3 (c). Comparison of average queue lengths as observed and estimated by Model III on lane 3.

In Figure 3 (a) – (c), the proposed model demonstrate the effectiveness of the estimation procedures for each lane, meanwhile Model III estimates the average queue lengths that is nearly identical to the observed average queue lengths per each cycle on individual lanes.

4. Concluding remarks

We apply a deep learning method to estimate cycle-based proportions of downstream arrivals used to predict downstream arrivals for the number of vertically queued vehicles in the current cycle. The key challenges are to construct the LSTM network. The primary contribution of the proposed method is to enhance the predictive accuracy of lane-based queue lengths in the future cycles using the historical patterns of proportional downstream arrivals. A major advantage of implementing an integrated deep learning process compared to the previously Kalman-filter-based queue estimation approach is that there is no need to calibrate the co-variance matrix and tune the gain values (parameters) of the estimator. The proposed method consists of discriminant models and a downstream arrivals estimation method using a comprehensive deep learning method to predict short-term queue lengths at each lane.

In the simulations, three models were examined, including the traditional conservation equation and the developed models that considered the discriminant models and the downstream arrivals as estimated by the Kalman filter and the LSTM network. It was found that the calibrated discriminant models and the LSTM network were effective. The proposed method performed well for the training set and for four different validation sets with 200m upstream detectors. The results of the computer simulations showed that the real-time estimations of lane-based queue lengths based on the discriminant models and downstream arrivals as estimated by the LSTM network outperformed the other methods. The results of the computer simulations demonstrated that the proposed method was robust and accurate for the estimation of lane-based queue lengths in real time under a wide range of traffic conditions. Estimating downstream arrivals with the LSTM network enhanced the accuracy of predictions by minimizing the error terms caused by lane-changing behaviour. In future research, these models can be used to establish group-based adaptive control logic in real time, or to adjust the duration, start time and sequence of signal groups by a few seconds or minutes for both isolated signalized junctions and area traffic control.

Acknowledgements

The work described in this manuscript was financially supported by a Postdoctoral Fellowship grant in connected and autonomous traffic systems within the Department of Civil and Natural Resources Engineering, College of Engineering at the University of Canterbury.

References

Akçelik, R., 1980. Time-dependent expressions for delay, stop rate and queue length at traffic signals. *Australian Road Research Board*, Melbourne, Australia.

Barcelo, J., Ferrer, J.L. and Montero, L., 1989. AIMSUN: Advanced Interactive Microscopic Simulator for Urban Networks. *Vol I: System Description, and*, *2*.

Comert, G., Cetin, M., 2009. Queue length estimation from probe vehicle location and the impacts of sample size. *European Journal of Operational Research* 197, 196-202.

Cheng, Y., X. Qin, J. Jin, B. Ran, and J. Anderson, 2011. Cycle by cycle queue length estimation for signalized intersections using sampled trajectory data. In 90th Annual Meeting of the Transportation Research Board. Washington, D.C.

Cheng, Y., X. Qin, J. Jin, and B. Ran, 2012. An exploratory shockwave approach to estimating queue length using probe trajectories. *Journal of Intelligent Transportation Systems 16*, 12-23.

Christofa, E., Argote, J. and Skabardonis, A., 2013. Arterial Queue Spillback Detection and Signal Control based on Connected Vehicle Technology. *Transportation Research Record: Journal of the Transportation Research Board 2356*, 61-70.

LeCun, Y., and Bengio, Y., 1995. Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks*, 3361(10).

Lee, S., Wong, S.C. and Li, Y.C., 2015. Real-time estimation of lane-based queue lengths at isolated signalized junctions. *Transportation Research Part C: Emerging Technologies*, 56, pp.1-17.

Lee, S., Wong, S. C., 2017a. Group-based approach to derivatives of delay models based on incremental queue accumulations for isolated signalized junctions. *Transportmetrica B: Transport Dynamics*, 1-21.

Lee, S., Wong, S. C., 2017b. Group-based approach to predictive delay model based on incremental queue accumulations for adaptive traffic control systems. *Transportation Research Part B: Methodological*, *98*, 1-20.

Lee, S., Wong, S. C., Varaiya, P., 2017a. Group-based hierarchical adaptive traffic-signal control part I: Formulation. *Transportation research part B: methodological*, *105*, 1-18.

Lee, S., Wong, S. C., Varaiya, P., 2017b. Group-based hierarchical adaptive traffic-signal control Part II: Implementation. *Transportation research part B: methodological*, *104*, 376-397.

Lindley, D.V., 1952. The theory of queues with a single server. *Mathematical Proceedings of the Cambridge Philosophical Society*. Cambridge Univ. Press, pp. 277-289.

Lighthill, M.J., Whitham, G.B., 1955. On kinematic waves. II. A theory of traffic flow on long crowded roads. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 229, 317-345.

Hochreiter, S., and Schmidhuber, J., 1997. "Long short-term memory." *Neural computation*, 9(8), 1735-1780.

Richards, P.I., 1956. Shock waves on the highway. *Operations Research* 4, 42-51.

Webster, F., 1958. Traffic signal settings. *Road Research Technical Paper* no. 39. Road Research Laboratory.